Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

IUST
Iran University of
Science and Technology

Development of a Digital Stethoscope for Enhancing Real-Time Respiratory Diagnostics

Julie Roslita Rusli1 *(C.A.), Muhamad Syahirin Danial Noor Shahrin*, Nurul Izzati Binti Che Abdu Patah*, Izanoordina Ahmad*, Siti Marwangi Mohamad Maharum* and Sairul Izwan Safie**.

Abstract: Digital stethoscopes represent a significant advancement in medical diagnostics, addressing the limitations of traditional auscultation methods, which often suffer from diagnostic delays and inefficient workflows. This digital stethoscope facilitates real-time diagnosis through machine learning and remote monitoring, utilizing the ESP32's ADC and Wi-Fi capabilities to wirelessly send audio data to a remote server for comprehensive analysis. By integrating modern technologies such as the ESP32 microcontroller and the MAX9814 microphone module, these devices capture and transmit high-fidelity respiratory sounds, overcoming the challenges of imprecision and time lag in conventional methods. Initial tests have demonstrated the device's ability to capture clear respiratory sounds, underscoring its potential for effective remote health monitoring and telemedicine. These improvements aim to enhance diagnostic accuracy, facilitate early diagnosis, and ultimately improve patient outcomes, showcasing the significant potential of digital stethoscopes to transform respiratory diagnostics and patient care, particularly in remote and telemedicine settings. In this research, a prototype of a digital stethoscope for respiratory diagnostics was developed and evaluated. The obtained results from the prototype measurements demonstrated that the proposed system could be a solid starting point for the actual implementation of an advanced respiratory monitoring system.

Keywords: Digital stethoscope, Respiratory diagnostics, Remote health monitoring

1 Introduction

RESPIRATORY diseases pose a major global health challenge, impacting millions of individuals worldwide. These conditions affect the lungs and other parts of the respiratory system, resulting in serious health complications. Accurate and timely diagnosis is crucial for effective treatment and management of these diseases [1]. The integration of digital technologies into healthcare represents a transformative shift in medical diagnostics and patient care, effectively tackling long-standing challenges in respiratory diagnostics such as delays and inefficiencies associated with traditional auscultation methods. Digital stethoscopes have emerged as a ground breaking innovation, harnessing the power of advanced microcontrollers and sensor technologies to enhance diagnostic accuracy and facilitate remote monitoring. With the increasing use of machine learning and neural network algorithms, these devices offer significant improvements in diagnostic capabilities and the potential for timely and effective intervention [2-3].

The development of a digital stethoscope specifically designed for respiratory diagnostics holds considerable promise for advancing patient care, particularly within telemedicine and remote healthcare contexts. This

Iranian Journal of Electrical & Electronic Engineering, 2025.
Paper first received 26 Dec. 2024 and accepted 22 Feb. 2025.
* The author is with the Department of Electronics Technology, British Malaysian Institute Campus, Universiti Kuala Lumpur, Gombak, Selangor, Malaysia.
E-mail: julie@unikl.edu.my, syahirir.shahrin@s.unikl.edu.my, izzati.abduulie@s.unikl.edu.my, izanoordina@unikl.edu.my , sitimarwangi@unikl.edu.my.
** The authors are with the Department of Plan Engineering Technology, Malaysia Institute of Industrial Technology, Universiti Kuala Lumpur, Johor Baharu, Johor, Malaysia.
E-mails: sairulizwan@unikl.edu.my.
Corresponding Author: julie@unikl.edu.my.

research aims to enable remote patient monitoring, allowing healthcare providers to assess respiratory health from a distance and supporting early detection and intervention for respiratory conditions. The integration of machine learning algorithms further boosts diagnostic accuracy and efficiency, contributing to improved patient outcomes and more proactive management of respiratory health.

A review of recent studies on digital stethoscopes highlights both advancements and limitations within the field. Sowrav Chowdhury et al. (2022)[4] explored self-assembled stethoscopes equipped with an electret condenser microphone and AI-based diagnostics. While their approach showed promise in terms of AI integration, it was hindered by the lack of signal amplification and real-time audio output for healthcare providers, limiting data presentation to waveform representation only. R. Chitra et al. (2023)[5] developed a digital stethoscope utilizing a NodeMCU ESP8266 and a CNN model, achieving an impressive 95% accuracy in lung sound monitoring. However, this system faced issues related to cloud dependency and the absence of real-time audio output, which could impact its practical utility. Muhammad Waqar et al. (2019)[6] designed a stethoscope based on Arduino DUE that provided high-quality audio and real-time graphical display but lacked diagnostic capabilities and digital data output, restricting its use to visual monitoring only. P.-W. Lo Frank et al. (2017)[7] developed a Bluetooth-powered wearable stethoscope using ATmega328P, offering long-distance monitoring and robust data storage capabilities but lacking diagnostic functionalities.

This research presents the development of a prototype digital stethoscope for respiratory diagnostics, utilizing the ESP32 microcontroller and MAX9814 microphone module, which effectively addresses diagnostic delays and precision issues inherent in traditional auscultation methods. By leveraging neural networks, the system accurately analyzes respiratory sounds and detects abnormalities, offering a robust solution for respiratory health monitoring and early intervention. To mitigate limitations related to real-time audio output and cloud dependency, the study proposes a dedicated web-based platform for real-time audio streaming and analysis, ensuring healthcare providers can access critical audio data efficiently. To overcome the limitations associated with real-time audio output due to cloud dependency, the research proposes the implementation of a dedicated web-based platform. This platform will facilitate real-time audio streaming and analysis, ensuring that healthcare providers can access critical audio data without relying solely on cloud services. Additionally, the study outlines future research directions and potential applications, highlighting the transformative potential of this digital stethoscope system for enhancing patient care and advancing respiratory diagnostics. By integrating these advancements, the digital stethoscope aims to improve diagnostic accuracy and efficiency in clinical settings.

2 Methodology

The research methodology comprises the process flow, block diagram, flowchart, and circuit diagram of the research. The process flow of the digital stethoscope research, as depicted in Fig. 1, consists of several key stages: Data Acquisition, Pre-processing, Data Training, Feature Extraction, and Output Classification. Each stage is integral to ensuring accurate and effective analysis of respiratory sounds for diagnostic purposes.

2.1 Data Acquisition

The process begins with data acquisition, where raw audio data is collected from the stethoscope. This stage involves recording the respiratory sounds using the digital stethoscope equipped with the MAX9814 microphone module. The high sensitivity and automatic gain control (AGC) of the MAX9814 ensure that even subtle respiratory sounds are captured accurately.

2.2 Pre-processing

Once the raw audio data is collected, it moves to the pre-processing stage. This crucial step involves several sub-processes:

- *Filtering*: Noise removal techniques are applied to eliminate any unwanted background noise that might have been captured along with the respiratory sounds.
- *Data Cleaning*: This involves removing any artifacts or irregularities in the data to ensure it is clean and reliable for further processing.
- *Data Type Conversion*: The audio data may need to be converted into a suitable format or type that can be processed by the subsequent stages.

2.3 Data Training

In the data training stage, the pre-processed data is split into different subsets for training, validation, and testing purposes. This step is essential for building and evaluating the neural network model. The data is split into the following proportions: 80% for the Training Set to train the neural network model, 10% for the Test Set to evaluate the model's performance and accuracy, and 10% for the Validation Set to fine-tune the model and prevent overfitting. These specific percentages are based on common practices in machine learning.

The model used in this research is a Convolutional Neural Network (CNN), which is well-suited for analyzing and classifying respiratory sounds. The CNN model consists of multiple convolutional layers that extract features from the input data, followed by pooling layers that reduce the dimensionality of the feature maps. The extracted features are then passed through fully connected layers that perform the final classification task. The model is trained using backpropagation, which adjusts the weights and biases of the network to minimize the loss between the predicted and actual outputs.

The 80% training set is used to train the CNN model, which is sufficient for the model to learn the patterns and relationships in the respiratory sound data. The 10% test set is used to evaluate the model's performance and accuracy, providing an unbiased estimate of the model's ability to generalize to new, unseen data. The 10% validation set is used to fine-tune the model's hyperparameters, such as the learning rate, batch size, and number of epochs, to prevent overfitting. By monitoring the validation set performance during training, the model can be optimized to achieve the best balance between training set performance and generalization to new data.

For audio data, the process involves transforming audio signals into a format suitable for CNNs, often by representing them as Mel-Frequency Cepstral Coefficients (MFCCs). MFCCs are features derived from the audio signal that represent the short-term power spectrum of a sound. They can be used as input to CNNs for audio classification tasks.

2.4 Feature Extraction

Feature extraction is a critical stage where meaningful features are extracted from the audio data to be used by the neural network for classification. In this research, Mel-Frequency Cepstral Coefficients (MFCC) are used to recognize the audio. The Mel-Frequency Cepstral Coefficients (MFCC) method is a widely adopted and prominent approach for extracting features from audio signals, and ongoing research is focused on improving its performance. In fact, the majority of speaker identification systems rely on MFCC as their primary feature extraction technique [8]. These coefficients represent the short-term power spectrum of the sound signal and are widely used technique in audio processing and speech recognition tasks [1]. They help in capturing the essential characteristics of the respiratory sounds.

2.5 Output Classification

The final stage is output classification, where the neural network classifies the respiratory sounds into different categories based on the extracted features. The classes for the outputs include:

- Healthy
- Chronic Obstructive Pulmonary Disease (COPD)
- Upper Respiratory Tract Infection (URTI)
- Lower Respiratory Tract Infection (LRTI)
- Asthma
- Bronchiectasis
- Pneumonia
- Bronchiolitis

In medical scenario, the goal is to detect and classify various diseases or conditions, with the healthy class serving as a baseline [9]. The classification results are displayed in the terminal window on a laptop, providing real-time diagnostic information. This information can be used by healthcare professionals to identify and diagnose respiratory conditions promptly and accurately. However, when a new input lies outside of the predefined classes, several things could happen. The neural network may misclassify the input into one of the existing classes, leading to incorrect diagnosis or treatment [10]. Alternatively, the network may output a probability distribution that is uncertain or ambiguous, indicating that the input does not fit neatly into any of the predefined classes [11]. If the network is trained with an out-of-distribution detection mechanism, it may recognize that the new input is outside of the training data distribution and flag it as an anomaly or unknown class [10].

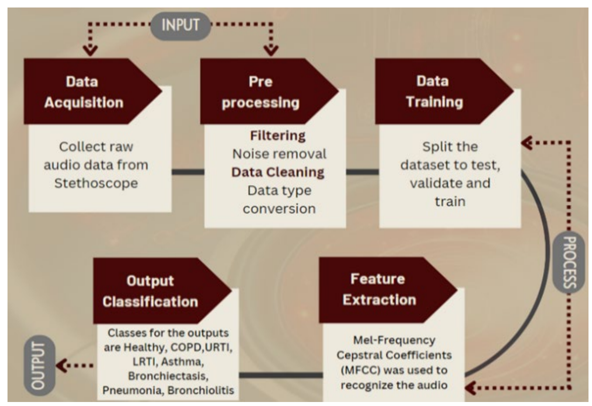


Fig 1. Process Flow of the Research

As shown in Fig. 1, the process flow outlines the proposed proof-of-concept development of a digital stethoscope for respiratory diagnostics using the ESP32 microcontroller and MAX9814 microphone module. The system captures respiratory sounds via the microphone module, which are then processed and transmitted by the ESP32 to a remote server for analysis. The digital stethoscope operates by using the MAX9814 microphone module to capture respiratory sounds, which are converted to digital data by the ESP32's analog-to-digital converter (ADC). This data is then sent wirelessly to a developed Flask remote server via Wi-Fi using the HTTP/HTTPS protocol for storage and further analysis. Flask is a web framework written in Python. It’s designed to be simple and flexible, allowing developers to build web applications with minimal setup and configuration. The server uses advanced algorithms to analyze respiratory sounds and detect any abnormalities.

2.6 System Architecture and Workflow

As shown in Fig. 2, the block diagram of the proposed digital stethoscope system includes three main components: the input stage, the processing stage, and the output stage. Each of these components plays a crucial role in capturing, processing, and displaying respiratory sounds for diagnostic purposes.

Input Stage:
The input stage begins with the stethoscope, which is used to capture respiratory sounds from patients. This device functions as an acoustic sensor, detecting the subtle sounds produced within the respiratory system. These sounds are then directed to the MAX9814 microphone module. The MAX9814 is a sophisticated electret microphone amplifier equipped with automatic gain control (AGC). The AGC is crucial for ensuring that even the faintest respiratory sounds are amplified and captured with high clarity. This module converts the acoustic signals into electrical signals that can be processed by the subsequent stage in the system.

Processing Stage:
At the core of the digital stethoscope system is the ESP32 microcontroller, which serves as the primary processing unit. The main advantages of the ESP32 is renowned for its powerful processing capabilities and versatile features, including an integrated Analog-to-Digital Converter (ADC). The electrical signals received from the MAX9814 are fed into the ESP32, where the ADC digitizes these analog signals into a format suitable for digital processing [12]. The ESP32 then processes these digital signals, performing tasks such as filtering and feature extraction. Additionally, the ESP32's built-in Wi-Fi capability enables it to transmit the raw data to a remote server for further analysis.

Output Stage:
The final stage of the system involves displaying the processed respiratory sound data on a laptop's terminal window. This terminal window acts as a real-time interface for monitoring and visualizing the captured sounds. The data is transmitted from the ESP32 to the laptop via Wi-Fi, where it is received and displayed in a format that allows even individuals with less knowledge of healthcare to observe the disease classification for the sound. This setup enables immediate access to the sound data, facilitating quick diagnostic assessments and continuous monitoring. By leveraging the capabilities of modern microcontrollers and advanced signal processing techniques, this digital stethoscope system enhances the accuracy and timeliness of respiratory diagnostics, providing a robust solution for patient care and early intervention.

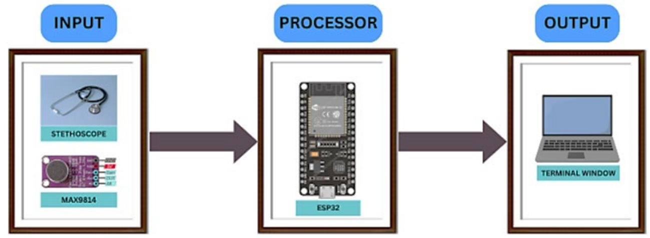


Fig 2. Block Diagram of the Hardware

As shown in Fig.3, the stethoscope will be connected to the microphone MAX9814 using a shrink tube to fit the mic into the stethoscope cord, and the pinout of the MAX9814 will be connected to the ESP32. Then the ESP32 will be connected to the desktop to run the code and server.

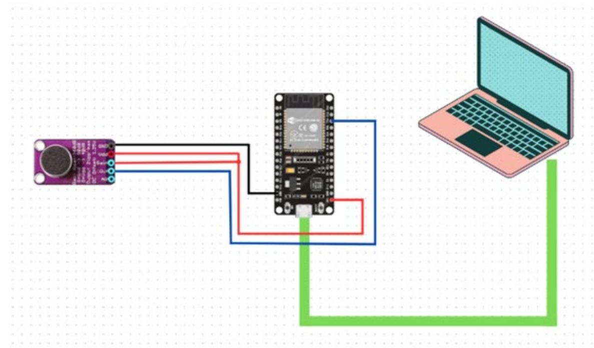


Fig 3. Circuit Diagram (Controller)

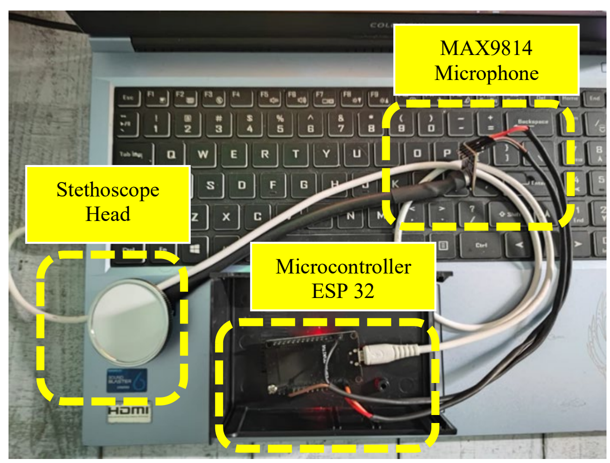


Fig 4. Complete Hardware of the Research

There are three sets of components in Fig. 4 which include a microphone amplifier module (MAX9814), a processor (ESP32), a desktop (laptop or PC), and the power comes from the desktop. Therefore, only one socket outlet is used. Any fault with the component can be detected using the desktop with different codes to test the components. This customized approach guarantees that any problems arising from either GPIO or power can be quickly located and fixed, improving the overall dependability.

2.7 Acquisition Method

The sequential or single-channel acquisition method involves using one microphone to capture audio data from different locations on the chest sequentially. This approach contrasts with multichannel acquisition, where multiple microphones record simultaneously from various locations [13]. In the single-channel method, the microphone is placed at a specific chest location, and the lung sounds are recorded continuously. If the sound is not properly recorded or unclear, the microphone is then moved to the next location, and the process is repeated. This method is more straightforward and cost-effective than simultaneous acquisition but requires careful handling to ensure consistent and comparable recordings from each location. Lung sounds are typically recorded from specific anatomical locations on the chest to capture the full range of respiratory sounds. The following chest locations were used in this research:

- *Trachea (Tc)*: The trachea is the large airway that conducts air from the larynx to the bronchi. Recording from the trachea captures breath sounds directly from the upper respiratory tract, which are generally louder and more turbulent.
- *Anterior Left (Al)*: This location is on the left front side of the chest, overlying the left lung. Sounds recorded here primarily capture the sounds generated in the left lung's upper lobe.
- *Anterior Right (Ar)*: Similar to the Anterior Left, this location is on the right front side of the chest, capturing sounds from the right lung's upper lobe.
- *Posterior Left (Pl)*: This location is on the back of the chest, over the left lung. Posterior recordings are essential for capturing sounds from the lower lobes of the lungs, which are crucial in diagnosing conditions like pneumonia.
- *Posterior Right (Pr)*: Located on the back of the chest over the right lung, this site captures sounds from the right lung's lower lobes.
- *Lateral Left (Ll)*: This location is on the left side of the chest. Lateral recordings help capture sounds from both the upper and lower lobes of the lungs, providing a comprehensive assessment of lung function.
- *Lateral Right (Lr)*: Similar to the Lateral Left, this location is on the right side of the chest and captures a broad range of lung sounds from the right lung.

2.8 Sampling Rate

For the dataset used in training, the sampling rate is set at 16,000 Hz (16 kHz)[14]. This decision was made to ensure consistency between the data used for training and the real-time data captured during the research's deployment. By using a uniform sampling rate, the model can better interpret and analyse the audio signals, leading to more accurate predictions and classifications of respiratory sounds. The 16 kHz sampling rate is well-suited to capturing the key frequency range of respiratory sounds, ensuring that the model is trained on data that reflects the actual conditions it will encounter in practice.

As shown in Fig. 5, the operation starts when the software program is executed on the ESP32 microcontroller. Once the system is powered on, the ESP32 will be connected to the Wifi to transmit the data, and the MAX9814 microphone module begins capturing respiratory sounds. The ESP32 processes these sounds and converts them from analog to digital data. If the sounds detected indicate any abnormalities, the data is transmitted to the remote server. The server, preloaded with the machine learning model, then analyzes the data, and the results are used to inform healthcare providers through a monitoring interface. This setup allows for real-time analysis and remote monitoring of respiratory health.

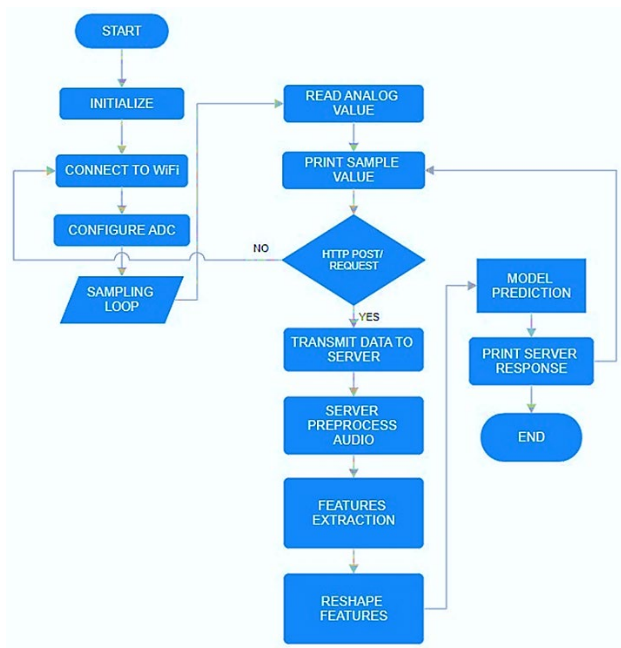


Fig 5. Flowchart of the Software Workflow

If the respiratory sounds are within normal parameters, the system continues to collect and transmit data to the server for continuous monitoring. This ensures that any potential issues can be detected early, and healthcare providers are alerted promptly via the monitoring interface, enhancing the efficiency and reliability of respiratory diagnostics [15]. In this way, the proposed digital stethoscope system aims to provide a more precise and timely diagnosis of respiratory conditions, leveraging modern microcontroller technology and advanced signal processing techniques. The integration with neural networks further enhances the system's capability to accurately analyze respiratory sounds and detect abnormalities, offering significant potential for improving patient care and facilitating early intervention.

3 Results and Discussion

This research was developed to establish a system that can accurately capture and analyse respiratory sounds for medical diagnostics. The primary objectives are to design a digital stethoscope integrated with a neural network capable of detecting abnormalities in respiratory sounds, to develop a working prototype for clinical testing and validation, and to implement an IoT application that enables real-time monitoring of respiratory health for healthcare providers [16]. Fig. 6 displays the training and validation loss against the number of epochs. The x-axis represents the epochs, which denote the number of complete passes through the entire training dataset. The y-axis measures the loss, a metric that quantifies the error between the model's predictions and the actual values.

The red line on the graph traces the training loss, while the blue line traces the validation loss. At the outset, both training and validation losses are high, indicating a significant error in the model's initial predictions. However, as the epochs progress, both losses decrease rapidly, showcasing the model's learning capability. The training loss continues to decrease and eventually stabilizes at a low value. This indicates that the model is effectively minimizing error on the training dataset, achieving a high degree of fit. Simultaneously, the validation loss follows a similar trend, decreasing and stabilizing at a point closely aligned with the training loss. This close alignment between training and validation loss suggests that the model generalizes well to unseen data, avoiding the common pitfall of overfitting [17].

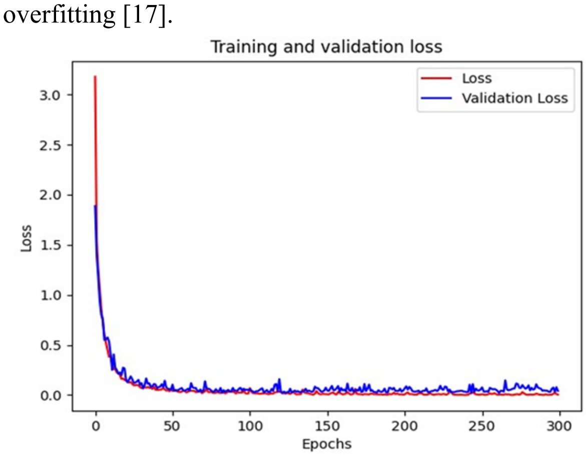


Fig 6. Training and Validation Loss for Machine Learning Model

Fig. 7 presents the training and validation accuracy, showing a consistent improvement in accuracy across the 300 epochs. The final accuracy reaches 0.99, demonstrating the model's high capability in correctly identifying respiratory sound patterns and detecting abnormalities. This high accuracy is crucial for ensuring reliable diagnostics and effective monitoring of respiratory health. The disparity in the accuracy and loss values between training and validation data can be attributed to the quality and quantity of the training data, as well as the optimization of neural network parameters. The high accuracy achieved suggests that the neural network model is well-suited for the task of analyzing respiratory sounds and providing accurate diagnostic information. The results indicate that the developed digital stethoscope system, integrated with a neural network, performs exceptionally well in analyzing respiratory sounds. This system holds significant promise for enhancing respiratory diagnostics, enabling early detection of abnormalities, and improving patient outcomes through timely intervention.

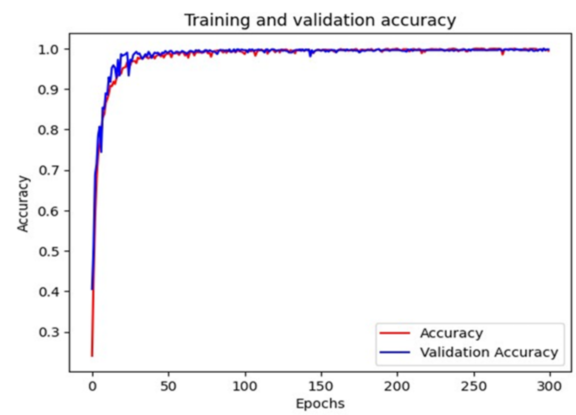


Fig 7. Training and Validation Accuracy for Machine Learning Model

Fig. 8 shows the confusion matrix for the neural network's predictions on the test data. The confusion matrix includes the following classes: asthma, bronchiectasis, bronchiolitis, COPD, healthy, LRTI, pneumonia, and URTI. The high accuracy observed in the confusion matrix results from up sampling, which was employed to counter the dataset's class imbalance. This technique ensures that the model is trained equally across all classes, thereby improving its ability to correctly classify each condition. The confusion matrix highlights the model's performance [18] in differentiating between various respiratory conditions, demonstrating high precision and recall across most classes. This is indicative of the model's robustness and its potential utility in clinical settings for accurately diagnosing respiratory conditions.

The provided ROC curve in Fig. 9 plot serves as a comprehensive evaluation of a multi-class classification model, detailing its proficiency in distinguishing between several classes: Asthma, Bronchiectasis, Bronchiolitis, COPD, Healthy, LRTI, Pneumonia, and URTI. This visualization employs several key elements to effectively communicate the model's diagnostic capabilities. The axes are fundamental to interpreting the plot. The x-axis represents the False Positive Rate (FPR), defined as the proportion of negative instances incorrectly classified as positive. Meanwhile, the y-axis denotes the True Positive Rate (TPR), also known as sensitivity or recall, which measures the proportion of actual positive instances correctly identified by the model. These metrics collectively provide insight into the model's accuracy and reliability.

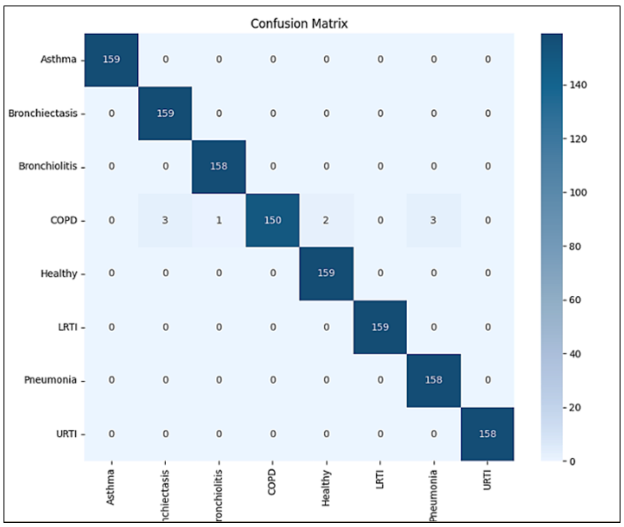


Fig 8. Confusion Matrix

At the heart of the plot is the Receiver Operating Characteristic (ROC) curve, a graphical representation illustrating the diagnostic ability of a classifier as its discrimination threshold varies. For multi-class classification problems, the ROC curve is plotted for each class versus all others using a one-vs-all approach. This method allows for a detailed evaluation of the model's ability to differentiate each specific class from the rest. A critical reference point in the plot is the dashed diagonal line, stretching from the bottom-left to the top-right, representing the performance of a random classifier. Any ROC curve positioned above this line indicates a model performing better than random guessing, providing a benchmark for comparison.

Each colored line on the plot represents the ROC curve for a different class. The accompanying legend provides the Area Under the Curve (AUC) for each class's ROC curve. The AUC is a singular value summarizing the classifier's overall performance, with an AUC of 1 indicating perfect classification. In this particular plot, each class achieves an AUC of 1.00. This perfect score signifies that the classifier can flawlessly distinguish each class from the rest, achieving perfect sensitivity and specificity. The ROC curves for all classes converge at the top-left corner of the plot, reflecting a 100% true positive rate (TPR = 1) without any false positives (FPR = 0) for each class.

The ROC curve plot demonstrates an exemplary performance by the multi-class classification model [19-20]. Each class's ROC curve aligns with the top-left corner, indicating flawless classification with perfect sensitivity and specificity. The AUC score of 1.00 for each class further underscores the model's impeccable ability to differentiate between various classes without errors. This ideal outcome, while rare in practical scenarios, suggests a highly effective and reliable classifier, capable of precise distinctions between different conditions and healthy states.

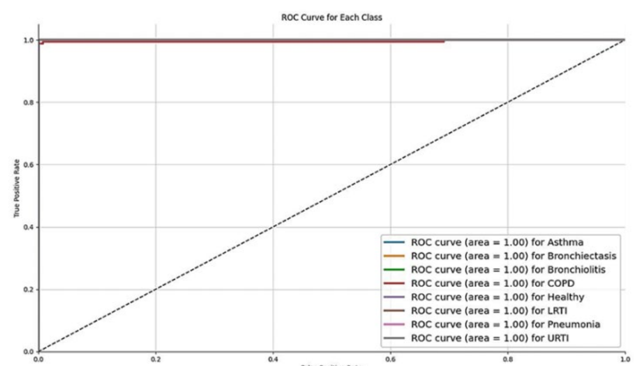


Fig 9. ROC curve (receiver operating characteristic curve)

Fig. 10 shows the output received in classification. The ESP32 will be connected to the internet first and print a response whether the Wifi is successfully connected. Then it will read the analog value and connect to the server via HTTP response to send the data. If the connection to the server fails, it will print "Failed to connect to the server. Please check the server URL and network connection." The terminal window then prints the response received from the server which is the output classification.

```
Sending data to server...
HTTP Response code: 200
Response: {"predicted_class":3,"predicted_label":"COPD"}
```

Fig 10. Output Received in Terminal Window

Fig. 11 demonstrates the successful data processing carried out by the Flask server. This figure confirms that data from the hardware has been successfully transmitted to the server. The

Flask server, which is preloaded with a machine learning model and audio processing capabilities, is designed to handle the newly received raw data efficiently. When the hardware sends data to the server, the server performs several crucial functions. First, it processes the raw data using audio processing algorithms. This step is essential for preparing the data for analysis by the machine learning model. The server then uses the machine learning model to analyze the processed data, giving out the output in the form of disease classification. The Flask server provides an IP address, which is a crucial feature for several reasons. By clicking on this IP address, users can verify that the server is still up and running. This ability to check the server's status ensures that the data processing pipeline remains operational, allowing for continuous monitoring and real-time data analysis.

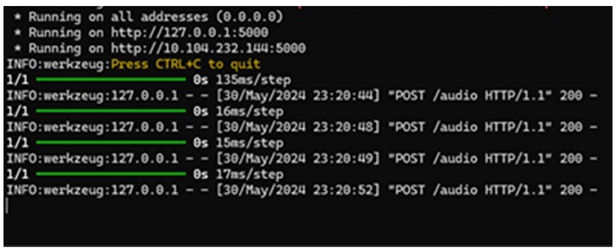


Fig 11. Data Log in the Server

As shown in Fig. 12, “Digital Stethoscope” was written on the web of the server. This confirmation indicates that the Flask server is up and running, accessible via the IP address provided in the command prompt during the server's initiation. Utilizing this Flask server offers significant advantages in handling data transmission. Instead of relying solely on the ESP32 microcontroller, which has limited memory capacity, the Flask server can manage and transmit larger volumes of data. The ESP32, while powerful and versatile for many embedded applications, may struggle with extensive data processing or storage due to its constrained memory resources.

By offloading the data transmission tasks to the Flask server, you can achieve enhanced data handling, allowing the server to manage larger datasets, perform more complex data processing, and store more information than the ESP32 could manage on its own. This setup offers scalability, as the Flask server can accommodate increasing amounts of data and more complex processing needs without being limited by the hardware constraints of the ESP32. Performance is also improved, as delegating data-intensive tasks to the server lets the ESP32 focus on its core functions, such as data collection and basic processing, leading to more efficient overall system performance.

Moreover, the Flask server can be accessed remotely through the provided IP address, enabling real-time monitoring, control, and data acquisition from anywhere with internet access. This setup also offers flexibility in development, as Flask, being a lightweight web framework, allows for easy development and integration of web-based interfaces, APIs, and data management tools, enhancing the functionality of the digital stethoscope system. In summary, deploying the Flask server to handle data transmission significantly enhances the capabilities of the digital stethoscope system, overcoming the memory limitations of the ESP32 and providing a robust platform for scalable and efficient data management.

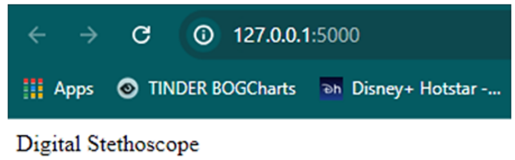


Fig 12. Text in the Web of the Flask Server

As long as the server is running, diagnostic processes can be carried out using the "curl" command in another command prompt window. Fig. 13 illustrates this process, demonstrating how the "curl" command can be utilized to diagnose a file located at a specified path. The "curl" command is a powerful tool used for transferring data to or from a server using various protocols. In this context, it facilitates communication with the Flask server, allowing users to send data for diagnosis and receive the results. To use the "curl" command for diagnosis, ensure that the file to be diagnosed is saved at the specified path on your system. This file contains the raw data that needs to be analyzed by the server. The "curl" command must be customized to match the file name and path. This customization involves specifying the correct file location and ensuring that the command points to the appropriate endpoint on the Flask server.

The process begins by invoking the "curl" command in a command prompt window. The command includes options and parameters that define the method of data transfer and the target URL, which is the Flask server's endpoint for diagnosis. By providing the path to the file and any necessary headers or data fields, the "curl" command sends the file to the server for processing. Once the server receives the file, it processes the data using its preloaded machine learning model and audio processing algorithms. The server then returns the results of the diagnosis, which can be viewed directly in the command prompt. This method allows for efficient and straightforward data submission and result retrieval without needing a web interface.

The “curl” command is a versatile tool that enables efficient communication with the Flask server for diagnostic purposes. Using the "curl" command for diagnosis offers several advantages. It provides a simple and scriptable way to interact with the server, making it easy to automate the diagnostic process or integrate it into larger workflows. Additionally, it allows for quick verification of the server's functionality and ensures that the diagnostic capabilities remain accessible as long as the server is running. The "curl" command enables efficient communication with the Flask server for diagnostic purposes. By customizing the command to include the correct file path and server endpoint, users can submit files for diagnosis and receive results directly in the command prompt. This approach simplifies the diagnostic process, making it accessible and easy to automate.

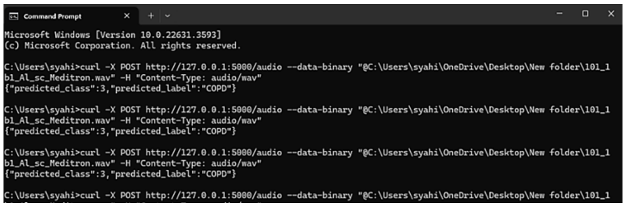


Fig 13. Diagnose the Audio File with Path Using Command Prompt

4. Conclusion & Recommendations

In this research, a prototype digital stethoscope for respiratory diagnostics was developed and evaluated. The prototype demonstrated strong potential as a foundational tool for implementing an advanced respiratory monitoring system. It effectively tackled critical issues such as diagnostic delays, treatment lags, and the limitations of traditional auscultation methods. The research's development holds promise for accurate analysis and identification of respiratory conditions, thereby reducing health risks and improving patient outcomes. By integrating machine learning algorithms and real-time monitoring capabilities, the system provides clear visibility into patients' respiratory health, facilitating early detection and timely intervention. However, future developments should focus on overcoming the current limitations of the system, including the need for quiet environments for audio recording, internet connectivity for server processing, and accurate placement of the stethoscope. The accurate analysis of lung sounds is crucial for diagnosing respiratory diseases, and advances in respiratory sound signal processing techniques have improved the detection of various conditions. Efforts should be made to collect a more accurate and balanced dataset to improve the system's performance and reduce class imbalance. In conclusion, the development and evaluation of the digital stethoscope prototype for respiratory diagnostics have shown that it effectively addresses key challenges such as diagnostic delays and the imprecision of traditional methods. By integrating advanced signal processing and machine learning algorithms, the system enables accurate and timely analysis of respiratory sounds. It facilitates real-time monitoring and continuous data transmission to a remote server, allowing for early detection of abnormalities and prompt healthcare provider alerts. This approach significantly improves diagnostic accuracy and patient outcomes, making it a promising solution for enhancing respiratory health monitoring and enabling timely interventions.

Acknowledgment

This research was conducted in of Universiti Kuala Lumpur British Malaysian Institute (UniKL BMI) and its publication is financially supported by the university. Therefore, the authors would like to thank UniKL BMI for the provision of laboratory facilities and financial support.

References

[1] D. M. Huang, J. Huang, K. Qiao, et al., "Deep learning-based lung sound analysis for intelligent stethoscope," Military Med. Res., vol. 10, no. 44, 2023. doi: 10.1186/s40779-023-00479-3.

[2] S. H. Lee, Y. S. Kim, M. K. Yeo, M. Mahmood, N. Zavanelli, C. Chung, J. Y. Heo, Y. Kim, S. S. Jung, and W. H. Yeo, "Fully portable continuous real-time auscultation with a soft wearable stethoscope designed for automated disease diagnosis," Sci. Adv., vol. 8, no. 21, May 2022.

[3] J. Zhao and S. Wang, "A Review of Respiratory Sound Signal Processing Techniques for Disease Detection," IEEE Access, vol. 8, pp. 155206-155222, 2020. doi: 10.1109/ACCESS.2020.3011850.

[4] S. Chowdhury, A. B. M. S. U. Doulah, and M. Rasheduzzaman, "Quality Assessment of Respiratory Sounds Extracted from Self-Assembled Digital Stethoscopes," in Proc. IEEE Conf. on Electrical and Electronics Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh, 2022, pp. 45-51.

[5] R. Chitra, N. Jayapreetha, D. Swetha, and S. Swetha, "Digital Stethoscope for Instant Monitoring for Cardiac Auscultation," in Proc. Int. Conf. on Electrical Engineering and Computer Science, Sri Sairam Engineering College, Chennai, India, 2023, pp. 123-130.

[6] M. Waqar, S. Inam, et al., "Arduino Based Cost-Effective Design and Development of a Digital Stethoscope," J. Med. Devices, vol. 13, no. 4, pp. 678-686, Oct. 2019.

[7] P.-W. F. Lo and M. Q.-H. Meng, "A Low-Cost Bluetooth Powered Wearable Digital Stethoscope for Cardiac Murmur," J. Healthc. Eng., vol. 8, no. 2, pp. 345-356, Jul. 2017.

[8] M. Saleh, N. Ibrahim, and D. Ramli, "Data reduction on MFCC features based on kernel PCA for speaker verification system," WALIA Journal, vol. 30, no. 2, pp. 56-62, 2014.

[9] G. M. M. Alshmrani, Q. Ni, R. Jiang, H. Pervaiz, and N. M. Elshennawy, "A Deep Learning

Architecture for Multi-Class Lung Diseases Classification Using Chest X-Ray (CXR) Images," Alexandria Eng. J., vol. 64, pp. 923-935, 2023.

[10] A. Nguyen, J. Yosinski, and J. Clune, "Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2015, pp. 427-436.

[11] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," in Adv. Neural Inf. Process. Syst., 2017, pp. 5574-5584.

[12] J. R. Rusli, S. Shafie, W. Z. W. Hassan, H. A. Majid, I. Ahmad, and M. A. Mustafa, "A Post-Silicon Validation Method for Low-Power 180 nm Dynamic Comparator in Differential 10-bit SAR ADC," in Proc. IEEE 9th Int. Conf. on Smart Instrumentation, Measurement and Applications (ICSIMA), Kuala Lumpur, Malaysia, 2023, pp. 199-204.

[13] S. Liang, Y. Li, and R. Srikant, "Enhancing the Reliability of Out-of-Distribution Image Detection in Neural Networks," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2018, pp. 8578-8586.

[14] X. Wang, Z. Zhang, and Y. Zhang, "Comparison of Sequential and Simultaneous Acoustic Measurement Techniques for Respiratory Sounds," J. Biomed. Eng., vol. 42, no. 4, pp. 1152-1163, 2019. doi: 10.1016/j.jbiomech.2019.03.014.

[15] S. Kwon, J. Park, and S. Lee, "Optimal Sampling Rate and Feature Extraction for Real-time Respiratory Sound Analysis," Sensors, vol. 21, no. 18, p. 6134, 2021.

[16] Y. W. Kuo, Y. C. Tsao, W. C. Chien, Y. M. Huang, and L. D. Liao, "Smart health monitoring and management system for organizations using radio-frequency identification (RFID) technology in hospitals or emergency applications," Emerg. Med. Int., vol. 2022, p. 2177548, 2022.

[17] L. B. Tolle, "Challenges in the Diagnosis and Management of Patients with Fibrosing Interstitial Lung Disease," Case Rep. Pulmonol., vol. 2022, p. 9942432, Feb. 2022. doi: 10.1155/2022/9942432.

[18] Y.-W. Ju, C. Hui, and Lun, "IoT-based wearable health monitoring device and its validation for potential critical and emergency applications," Front. Public Health, vol. 11, 2023.

[19] M. S. Kamel, J. L. Davidson, and M. S. Verma, "Strategies for Bovine Respiratory Disease (BRD) Diagnosis and Prognosis: A Comprehensive Overview," Animals, vol. 14, no. 4, p. 627, 2024.

[20] R. Wang and M. S. Bonney, "Novel Data Acquisition Utilizing a Flask Python Digital Twin Operational Platform," in Special Topics in Structural Dynamics & Experimental Techniques, M. Allen, S. Davaria, and R. B. Davis, Eds., Cham: Springer, 2023.

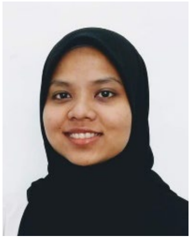
Biographies

Julie Roslita Rusli (SMIEEE) received her Master of Science in Electronics Eng. (Memory Testing) and Ph.D. in Electronics Engineering (Mixed Signal IC Design) from Universiti Putra Malaysia (UPM). She was a Validation Hardware Engineer in the Board Design Center at Intel Malaysia (M) Sdn. Bhd. from 2002 to 2004, where she was involved in signal integrity validation for newly designed motherboards.

From 2014 to 2015, she was involved in designing and implementing an ultra-low-power SAR ADC with Mimos Malaysia Sdn. Bhd. Julie Roslita is currently a senior lecturer at the Universiti Kuala Lumpur British Malaysian Institute. Her research interests include mixed-signal IC design, analog IC design, failure analysis, and IoT.

She is a senior member of IEEE and actively contributes to various IEEE societies, currently serving as Secretary of the IEEE Circuits and Systems Society (CAS) Malaysia Chapter for 2024. She is also a Professional Technologist certified by the Malaysian Board of Technologists. Her current research focuses on designing an ultra-low-power dynamic comparator for a differential 10-bit Successive Approximation Register (SAR) Analog-to-Digital Converter (ADC).

Muhamad Syahirin Danial Noor Shahrin recieved Bachelor of Electronic Engineering Technology with Honours at Universiti Kuala Lumpur British Malaysian Institutea. His research focuses on artificial intelligence applications in medical technology. He was as Technician at Sony EMCS in 2021, showcasing expertise in equipment maintenance and troubleshooting. Excelled in team collaboration, significantly enhancing operational efficiency. Demonstrated a strong commitment to safety and quality standards, with a knack for applying academic knowledge to practical scenarios, ensuring optimal equipment performance and reliability



Nurul Izzati Binti Che Abdu Patah is an undergraduate student pursuing a Bachelor of Electronic Engineering Technology (Medical Electronics) with Honours at Universiti Kuala Lumpur British Malaysian Institute (UniKL BMI), Malaysia. Her research focuses on artificial intelligence applications in medical technology, with her current project exploring "Deep Learning-Based Classification of Anesthesia Depth Using EEG and Bispectral Index." She is a Dean's List recipient for multiple semesters and is passionate about developing AI-driven solutions to enhance patient care and medical outcomes.

Izanoordina Ahmad received her Master of Science in Mechatronics (Singal and System) in 2008 from IIUM and Ph.D. in 2018 from University of Queensland. She was a R&D Engineer in Sharp Sdn. Bhd from 1998 to 1999, where she was involved in design and testing for newly designed DVDs. She is a Professional Technologist certified by the Malaysian Board of Technologists. Izanoordina is currently a senior lecturer at the Universiti Kuala Lumpur British Malaysian Institute. Her current research focuses on designing an WSN, IOT and AI.

Siti Marwangi Mohamad Maharum is a Senior Lecturer and currently responsible as the Head of Research & Innovation Section at Universiti Kuala Lumpur British Malaysian Institute. She completed her Bachelor of Engineering in Electrical-Telecommunications and PhD in Electrical Engineering from Universiti Teknologi Malaysia (UTM) in 2010 and 2015, respectively.

Her expertise spans across various Electrical & Electronic Engineering areas such as Microprocessor & Microcontroller, Millimeter Wave (mmWave), Interference Mitigation /Management, Internet of Things (IoT), Crosstalk in Analog IC Design, Artificial Intelligent (AI), and Machine Learning (ML).

Sairul Izwan Shafie received a Bachelor of Engineering (Electrical) and a Master of Electrical Engineering (Power) from the Universiti Teknologi Malaysia. He earned his Ph.D. in Electronic and Electrical Engineering from the University of Strathclyde, Glasgow, United Kingdom. With a wealth of knowledge and expertise in his field, he currently assumes the role of Associate Professor at the Universiti Kuala Lumpur Malaysian Institute of Industrial Technology.

His scholarly pursuits have garnered significant recognition, making him a respected figure within both the academic and research communities. His research pursuits are both diverse and impactful, covering an array of cutting-edge areas. His primary focus lies at the intersection of artificial intelligence, deep learning, and big data, where he endeavors to unravel the potential of these innovative technologies. Notably, his work delves into the realm of signal and image processing, where he explores novel approaches and applications for biometric, psychological, and physiological signals, images, and videos.